

Campbell Systematic Reviews

2008:12

First published: 19 August, 2008

Last updated: 29 August, 2008

Court-Mandated Interventions for Individuals Convicted of Domestic Violence

Lynette Feder, David B. Wilson, Sabrina Austin



THE CAMPBELL COLLABORATION

Colophon

Title	Court-mandated interventions for individuals convicted of domestic violence
Institution	The Campbell Collaboration
Authors	Feder, Lynette, Ph.D. Wilson, David B, Ph.D. Austin, Sabrina, M.S.
DOI	10.4073/csr.2008.12
No. of pages	46
Last updated	19 August, 2008

Citation	Feder L, Wilson DB, Austin S. Court-mandated interventions for individuals convicted of domestic violence. Campbell Systematic Reviews 2008:12 DOI: 10.4073/csr.2008.12
Copyright	© Feder et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Keywords	

Contributions	Not stated.
Support/Funding	Smith-Richardson Foundation
Potential Conflicts of Interest	The first listed author (Lynette Feder) was the primary investigator of an experiment assessing the effectiveness of a court-mandated counseling program conducted in Broward County, Florida. To best counter the potential conflict of interest, the review was made as transparent as possible and included a collaborator who had not been involved in any of the prior research reviewed here.

Corresponding author	Lynette Feder Liberal Arts & Sciences Portland State University PO Box 751 Portland, Oregon 97207-0751 USA E-mail: lfeder@pdx.edu
-----------------------------	---

Campbell Systematic Reviews

Editors-in-Chief Mark W. Lipsey, Vanderbilt University, USA
Arild Bjørndal, Norwegian Knowledge Centre for the Health Services &
University of Oslo, Norway

Editors

Crime and Justice David B. Wilson, George Mason University, USA

Education Chad Nye, University of Central Florida, USA
Ralf Schlosser, Northeastern University, USA

Social Welfare Julia Littell, Bryn Mawr College, USA
Geraldine Macdonald, Queen's University, UK & Cochrane Developmental,
Psychosocial and Learning Problems Group

Managing Editor Karianne Thune Hammerstrøm, The Campbell Collaboration

Editorial Board

Crime and Justice David Weisburd, Hebrew University, Israel & George Mason University, USA
Peter Grabosky, Australian National University, Australia

Education Carole Torgerson, University of York, UK

Social Welfare Aron Shlonsky, University of Toronto, Canada

Methods Therese Pigott, Loyola University, USA
Peter Tugwell, University of Ottawa, Canada

The Campbell Collaboration (C2) was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. C2 offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of C2's editors, librarians, methodologists and external peer-reviewers contribute.

The Campbell Collaboration
P.O. Box 7004 St. Olavs plass
0130 Oslo, Norway
www.campbellcollaboration.org

**COURT-MANDATED INTERVENTIONS FOR
INDIVIDUALS CONVICTED OF
DOMESTIC VIOLENCE
A Campbell Collaboration Systematic Review**

Lynette Feder, Ph.D.
Associate Professor

Liberal Arts & Sciences
Portland State University
PO Box 751
Portland, Oregon 97207-0751
USA
lfeder@pdx.edu

David B. Wilson, Ph.D.
Associate Professor
George Mason University
10900 University Blvd., MS 4F4
Manassas, VA 20110
USA
dwilsonb@gmu.edu

Sabrina Austin, M.S.
Criminology, Law & Society
University of California, Irvine
Irvine, CA 92697-7080
USA
seaustin@uci.edu

Abstract

Background: Survey research and analysis of police records, hospital emergency rooms and women's shelters have clearly established the severity of the domestic violence problem and the need to find programs to address this issue. Today, court-mandated batterer intervention programs (BIPs) are being implemented throughout the United States as one of the leading methods to address this problem. These programs emerged from the women's shelter movement and therefore contained a strong feminist orientation. They developed as group-based programs, typically using psychoeducational methods. Their aim was to get men to take responsibility for their sexist beliefs and stop abusing their partners by teaching them alternative responses for handling their anger.

Objectives: The aim of this systematic review is to assess the effects of post-arrest court-mandated interventions (including pre-trial diversion programs) for domestic violence offenders that target, in part or exclusively, batterers with the aim of reducing their future likelihood of re-assaulting above and beyond what would have been expected by routine legal procedures.

Search Strategies: We searched numerous computerized databases and websites, bibliographies of published reviews of related literature and scrutiny of annotated bibliographies of related literature. Our goal was to identify all published and unpublished literature that met our selection criteria.

Selection Criteria: We included experimental or rigorous quasi-experimental evaluations of court-mandated batterer intervention programs that measured official or victim reports of future domestic violent behavior. Rigorous quasi-experimental designs were defined as those that either used matching or statistical controls to improve the comparability of the groups. Given their importance in the literature, we also included rigorous quasi-experimental designs that used a treatment drop-out comparison.

Data Collection and Analysis: We coded characteristics of the treatment, sample, outcomes, and research methods. Findings were extracted in the form of an effect size and effect sizes were analyzed using the inverse-variance method. Official report and victim report outcomes were analyzed separately as were the different design types (i.e., random, quasi-experimental with a no treatment comparison, and quasi-experimental with a treatment dropout comparison).

Main Results: The mean effect for official reports of domestic violence from experimental studies showed modest benefit whereas the mean effect for victim reported outcomes was zero. Quasi-experimental studies using a no-treatment comparison had inconsistent findings indicating an overall small harmful effect. In contrast, quasi-experimental studies using a treatment dropout design showed a large, positive mean effect on domestic violence outcomes. The latter studies suffer, we believe, from selection bias.

Reviewer's Conclusions: The findings, we believe, raise doubts about the effectiveness of court-mandated batterer intervention programs in reducing re-assault among men convicted of misdemeanor domestic violence.

Overview

Domestic violence is defined as assaultive behavior involving adults who are married, cohabitating, or who have an ongoing or prior intimate relationship (Goolkasian, 1986). Research indicates just how pervasive this problem is today. Based upon crimes reported to the police in 1998, intimate partner homicides accounted for about 11% of all murders nationwide (Rennison & Welchans, 2000). The National Crime Victimization Survey (NCVS) indicated that there were about 1 million violent crimes committed against persons by their current or former spouses or significant others in 1998, with the vast majority (85%) being against female victims (Rennison & Welchans, 2000). These numbers speak to the importance of finding programs that can successfully intervene with domestic violence offenders.

Individual studies evaluating court-mandated batterer intervention programs (BIPs) have provided very mixed findings on their effectiveness. While this first wave of evaluation research consistently indicated high rates of success, their findings probably reflected the methodological shortcomings of the research rather than the programs' actual effectiveness in reducing violence. That was followed by a period in which more rigorous research was conducted. Unlike the earlier studies, these studies produced mixed results regarding the effectiveness of mandated batterer intervention programs in reducing violence. These mixed results possibly reflected differences in the rigor of the research methodology used to evaluate these programs along with differences in outcome measures utilized, length of time followed, and the integrity with which the intervention was implemented apart from additional programs and services that may have been provided at these different sites.

To date, two meta-analyses have been conducted studying the effectiveness of court-mandated counseling in reducing future violence among domestic violence offenders. Davis and Taylor (1999) included five quasi-experimental studies using a non-equivalent matched group design (they discarded one study because its results were viewed as anomalous) and two experimental studies with random assignment. They concluded that, “among the handful of quasi- and true experiments there is fairly consistent evidence that treatment works and that the effect of treatment is substantial” (Davis & Taylor, 1999, p. 69). Their analysis found a mean effect size (*d*) of 0.412 for experimental studies and 0.416 for quasi-experimental studies. In all, Davis and Taylor found a fairly substantial effect size for this intervention.

Babcock, Green & Robie (2003) examined a larger number of evaluations in their systematic review and meta-analysis. Their search yielded 17 quasi-experimental studies (where treatment completers were compared to treatment dropouts, no-shows and/or treatment rejects or to a matched comparison group that did not receive treatment) and 5 experimental designs (with random assignment to treatment and control conditions). Babcock and her colleagues concluded that “the effect size due to group battering intervention on recidivism of domestic violence is in the 'small' range” (Babcock, et al., 2003, p. 1043). “To a clinician, this means that a woman is 5% less likely to be re-

assaulted by a man who was arrested, sanctioned, and went to a batterers' program than by a man who was simply arrested and sanctioned" (Babcock, et al., 2004, p. 1004).

Babcock et al.'s study was a thorough and systematic review of the research literature. However, their inclusion of all quasi-experimental studies (including those failing to establish pre-intervention equivalency) combined with their failure to separately analyze effect sizes for these different types of quasi-experimental studies may lead towards a bias in favor of finding positive results if rigor of research design is in fact related to likelihood of finding treatment effectiveness (Feder & Forde, 2000; Weisburd, Lum & Petrosino, 2001).

Our systematic review uses meta-analytic procedures to synthesize the extant empirical evidence on the effects that court-mandated batterer intervention programs (including pre-trial diversion programs) have, over and above the effect of routine legal interventions, on rates of recidivism based upon rigorous research that has been conducted to date. Like Babcock and her associates, we conducted a systematic review to locate all studies conducted in the United States and elsewhere, whether published or not published. Also like Babcock, we included all experimental designs meeting our inclusion criteria. Unlike Babcock, we did not include all quasi-experimental studies but instead limited inclusion to those that address the problem of selection bias either via a match group design or statistical controls. Additionally, we provided separate analyses for each of the specific type of research designs so as to decipher the effect that this factor has on finding treatment effectiveness.

Background to Court-Mandated Batterer Programs

The idea of counseling male domestic violence offenders developed directly out of the women's shelter movement where advocates, working with battered women, realized that the only way to stop the cycle of violence was to change the behavior of the abuser (Feazell, Mayers and Deschner, 1984). It is not surprising, therefore, that these programs borrowed heavily from a feminist orientation. Typically, the various programs encouraged men to confront their sexist beliefs and accept responsibility for their past abuse, while teaching them alternative behaviors and reactions (e.g., anger management, assertiveness, relaxation techniques and communication skills).

The greatest growth in these different BIPs was brought about by the rise in pro-arrest domestic violence laws in the late 1980s. As police increased their rates of arrest for these offenses, pressure was placed on courts to deal with these offenders. Given this population's high rates of attrition from treatment programs, court-mandated BIPs were viewed as one method to ensure greater compliance while simultaneously serving as an alternative to over-crowded jails.

Soon after these court-mandated programs began appearing, studies evaluating their efficacy began surfacing. In this first wave of evaluation research, the results indicated suspiciously high rates of success in reducing the frequency and/or severity of subsequent violence amongst this offender population. A number of researchers noted that these findings may have reflected the methodological shortcomings inherent in these studies rather than the programs' actual efficacy in reducing violence (Chen, Bersani, Myers and Denton, 1989; Ford and Regoli, 1993; Gondolf, 1987).

More recent and rigorous research has resulted in inconsistent conclusions about the programs' effectiveness in reducing recidivism amongst domestic violence offenders. Four quasi-experimental matched comparison group studies found mixed results in terms of statistical significance regarding the efficacy of court-mandated treatments. Two studies, comparing treatment completers with treatment dropouts while matching for demographics and criminal histories, found men court-mandated into counseling to be significantly less likely to recidivate (Dutton, 1986; Gondolf, 1998). Another quasi-experimental study found no difference in future violence between those mandated into counseling and a comparison no-treatment group (Chen et al., 1989). In contrast, a quasi-experimental study by Harrell (1991) found that those mandated into counseling had significantly higher rates of recidivism than a comparable group of non-mandated domestic violence offenders.

These inconsistent results have also been found in the five experimental tests of court-mandated interventions conducted. One study found that men randomly assigned into a court-mandated counseling program were less likely to recidivate than those assigned into a no-treatment control group (Palmer, Brown and Barrera, 1992). Another found that men mandated into a 26 week counseling program had significantly lower rates of recidivism than those mandated into an 8 week counseling program or the control condition - 40 hours of community service program (Davis, Taylor and Maxwell, 1999). However, three experiments found no difference in men mandated into a counseling program than those given probation only (Ford and Regoli, 1993; Dunford, 2000; Feder and Duggan, 2001) or mandated into a conjoint therapy program, rigorous monitoring, or a group receiving no program or monitoring (Dunford, 2000).

As more communities are called upon to develop a coordinated response to the problem of domestic violence we will most likely see a continued increase in the number of court-mandated interventions. Understanding these programs' effectiveness in reducing future violence, therefore, becomes increasingly important.

Objectives

In 1984, the Attorney General's Task Force on Family Violence recommended court-mandated treatment as an addition to legal alternatives (U.S. Attorney General's Task Force on Family Violence, 1984). Yet twenty years later, the field remains uncertain about whether these programs are more effective in reducing future violence than legal interventions alone (e.g., arrest, prosecution, conviction and short jail stay and/or probation). The National Academy of Sciences has noted that "the urgency and magnitude of the problem of family violence have caused policy makers, service providers, and advocates to take action in the absence of scientific knowledge that could inform policy and practice" (Chalk & King, 1998, p. 2). Therefore, the aim of this systematic review is to assess the effects of post-arrest court-mandated interventions (including pre-trial diversion programs) for domestic violence offenders that target, in part or exclusively, batterers with the aim of reducing their future likelihood of re-assaulting above and beyond what would have been expected by routine legal procedures. Additionally, by investigating results by specific types of research designs implemented, this study will investigate the effect that methodological design has on outcome findings.

Methods

Criteria for Inclusion in the Systematic Review

Types of studies: Only studies using an experimental or rigorous quasi-experimental design were included. Experimental designs were defined as those using random assignment to treatment and control group(s). Rigorous quasi-experimental designs were operationalized as those addressing selection bias in the experimental and control group(s) through the use of multivariate statistical methods or a matched subject research design. For both experimental and quasi-experimental designs, control conditions could be no-treatment, or treatment as usual. That is, the no-treatment control condition could include routine legal interventions such as probation or a short jail stay. We excluded, however, referral to counseling or alternative programs designed specifically to reduce domestic violence (beyond any deterrent effect of jail or probation). We included quasi-experimental designs that used treatment dropouts as the control condition if the study statistically adjusted for baseline differences. These treatment dropout studies were analyzed separately and were included because of their importance in the literature. As discussed in the findings and discussion sections, we believe that the statistical adjustments for selection bias used in these studies was inadequate, upwardly biasing the findings.

Types of interventions: The intervention involved a post-arrest court-mandated intervention that, in part or exclusively, was aimed at the batterer and had as its goal decreasing the batterers' future likelihood of re-assaulting that or other partners. As so defined, pre-trial diversion programs were eligible for inclusion.

Types of participants: Only studies that used adult participants (operationalized as persons aged 18 years or older) of heterosexual intimate domestic violence, whether presently or formerly married, separated, divorced, cohabiting or dating were included in the meta-analysis. As long as the study included individuals who met this criteria it was included in the systematic review even if the study sample included others who fell outside these criteria.

Types of outcome measures: In order for a study to have been included in this systematic review it had to use an outcome measure of repeat domestic violence obtained at least six months post-treatment. This was defined as six months from the time that the treatment ended, that is, the individual completed his court-mandate. This criterion was based on Dunford's findings that evaluation studies collecting outcome data at the end of treatment were more likely to find effectiveness than those measuring outcomes for some period post-treatment (Dunford, 2000). This suggested that evaluations that were based solely on end-of-treatment assessments should be viewed cautiously. Additionally, to be included, a study had to include at least one outcome measure on repeat violence to that or other victims that used something more than offenders' self-reported repeat violence. As such, studies that included victim reports of the offender's abusive behavior or official measures of recidivism including arrest, charges or convictions were eligible for inclusion.

It needs to be noted that studies which exclusively relied on attitudinal changes were not included in this meta-analysis. Undoubtedly, any positive effects of these programs

would be mediated by other changes, such as attitudes and the acquisition of anger management strategies. Changes in these intermediate outcomes would be encouraging and these changes might lead to benefits not detected in the outcomes examined. However, the primary purpose of these programs is a reduction in partner abuse, hence our focus on this critical outcome. Additionally, attitudinal changes would rely on batterers' self-reports. Whether it is due to social desirability or to other unknown factors, more than a few researchers working in this field have found reason to doubt these accounts (Edleson & Brygger, 1995; Feder & Duggan, 2002; Tolman & Edleson, 1995). As such, the decision was made to limit outcomes to measures of continued abuse.

Sufficiency of data included: Finally, to be included the study needed to have reported sufficient data to permit computation of an effect size.

Search Strategy for Identification of Studies

Our goal was to identify and include all published and unpublished studies conducted in the United States or elsewhere from 1986 through January 2003 that met our inclusion criteria. Toward this aim, the first author (Lynette Feder), who had worked in this field for many years, canvassed a number of other researchers for additional studies, published or not, on the effectiveness of batterer intervention programs. The research team also searched computerized databases and websites, bibliographies of published reviews of related literature and scrutiny of annotated bibliographies of related literature (see below). The list is grouped in terms of those focused on: (1) published materials; (2) non-published materials; (3) governmental publications; and (4) existing registers of studies on domestic violence. It must be noted that some of the databases could be listed in multiple groups. That is, contained under "Published Materials" is Sociological Abstracts, Educational Resources Information Clearinghouse (ERIC), Criminal Justice Abstracts and others that contain unpublished as well as published literature with some containing international as well as national studies. Searches were conducted using the following databases and websites:

(1) Published Materials

- PsycINFO
- ERIC
- MEDLINE
- Sociological Abstracts
- Social Science Citation Index
- Lexis Nexis Legal
- Lexis Nexis Medical
- Social Work Abstracts
- Criminal Justice Abstracts

(2) Non-Published Materials

- Dissertation Abstracts International

(3) Governmental

- GPO Monthly Catalog (MOCAT)
- National Criminal Justice Research Service
- UK National Health Service NRR (National Research Register)

(4) Existing Registers or Studies on Domestic Violence

Social, Psychological, Criminological and Educational Trials Register
(C2-SPECTR)

PsiTri database of randomized and controlled trials in mental health

Babcock and Taillade, 1999

Davis & Taylor, 1999

Babcock, Green & Robie, 2003

Terms Used to Search: We used twenty-five keywords in three clusters to search for all experimental and quasi-experimental studies conducted on the effectiveness of court-mandated interventions for domestic violence offenders. Whenever appropriate a “wildcard” was used so as to search for the root of the word allowing for other possible derivations. (So, for instance, the term “eval*” was used to pick up evaluation, evaluate, evaluating, etc.) Cluster One related to the subject matter. Cluster Two sought to find citations using program or evaluation research keywords. Finally, Cluster Three used keywords related to outcomes. Terms within a cluster were connected with the Boolean “or” (i.e., an abstract with any one of the terms got selected) and the clusters were then connected with the Boolean “and” (i.e., an abstract with at least one of the terms in each cluster got selected). To make the resulting list more manageable, the search was restricted to titles and abstracts. If the title or abstract looked promising, the entire study was pulled and reviewed. The keywords within each cluster were:

Cluster One – Subject Words

Anger management

Batter(er/s)

Domestic assault

Domestic violence

Family violence

Spous(e/al) abuse

Physical abuse

Minneapolis Model

Duluth

Intimate partner violence

Cluster Two – Program Words

Defer(ral/ring/rred)

Program(s)

Treatment(s)

Intervention(s)

Diversion(ary)

Prosecu(te/tion/torial

Cluster Three – Outcome Words

Effect(s/ive/iveness)

Research(es)

Outcome(s)

Eval(uation/luations/ating)

Experiment(al)

Quasi(-experimental)

Random(ly)

Compar(ison/ing)

The graduate research assistant and the first author reviewed the titles and abstracted those that were identified through the search process. Studies that appeared likely to be eligible were retrieved in their entirety. Where disagreements occurred, the second author (David Wilson) was consulted and differences were resolved. The graduate assistant and first author were also responsible for reviewing the full text of all studies retrieved in their entirety to determine final eligibility for the meta-analysis. Again, where disagreements or uncertainties regarding the inclusion of a study arose, the second author's opinion was sought to resolve the decision.

Description of the Methods Used in the Component Studies

The methods used in the component studies were implied in the inclusion criteria. Specifically, the studies used a comparison group design with random assignment to conditions or compared naturally occurring program participants with similar individuals not participating in the program (e.g., domestic violence probationers). The domestic violence program was compared to a no treatment condition, with no treatment being defined as routine processing by the criminal justice system. The typical study assigned offenders or offenders assigned themselves (as in the quasi-experimental studies using drop-outs and no-shows) to the experimental (treatment) or control (routine) conditions following conviction or through a diversionary process that avoided conviction. The studies measured post-program rates of re-offending, that is, new instances of domestic violence to that or new partners. These measures may have included official arrest or convictions or reports from the abusers' victim (e.g., (ex-)spouses or (ex-)girlfriends).

Criteria for Determination of Independent Findings

To avoid the "double counting" of findings, two strategies were employed. First, multiple publications of an evaluation were treated as a single study in the synthesis. Second, multiple findings from a single study were categorized by outcome construct (i.e., official report and victim report) and only a single effect per construct was used in any analysis. For the official report effect sizes, a decision rule for determining which effect to use in an analysis if multiple effects were available was used that was independent of the size of the observed effect. More specifically, preference was given to measures of arrest over conviction and estimates that adjusted for baseline features over non-adjusted estimates. Additionally, effect sizes reported for a longer time-frame (e.g., 12-months instead of 6-months) were selected over those of a shorter time-frame. The logic was to select an official report that was as close to the behavior (domestic abuse) as possible. As such, arrests involve fewer decisions on the part of the criminal justice system than do convictions. For victim report measures, all effect sizes measuring domestic violence were averaged and the composite used in the analyses. As with official reports, effect sizes for a longer time-frame or follow-up period were selected and averaged, excluding the same constructs measured at an earlier time point.

Details of Study Coding Categories

The specific items coded and the categories of those items were developed as the research team interacts with the literature. Data were coded to reflect characteristics of the

treatment programs, participants, and research methods. In addition, all outcomes of interest were coded as an effect size along with related information. Items were developed for each of these areas. A list of constructs coded is as follows.

(1) Treatment: type of treatment, participant dropout from treatment, treatment integrity, length of treatment, treatment setting, treatment provider, treatment philosophy.

(2) Participants: representativeness of sample, age, geographic location.

(3) Research methods: nature of the assignment to conditions, integrity of the assignment process, study level attrition, differential attrition between conditions, use of statistical controls, use of matching.

(4) Effect size: data necessary for computation of the effect size (sample sizes, proportions, frequencies, etc.), nature of the outcome measure, source of the outcome measure (victim reports and/or police records), time frame for the outcome measure.

The coding protocol allowed for the coding of multiple effect sizes per study (see section above for description of the methods for handling the dependencies this produces). Coding was performed on paper coding sheets similar to a paper survey form. The data were entered into a computer datafile for analysis and storage. The coding forms are provided in Appendix Two.

Coded study characteristics served primarily a descriptive function, detailing the nature of the studies, and were used in moderator analyses of effect sizes. Because the number of eligible studies was small, moderator analyses were not performed. However, effect sizes for the basic research design categories (i.e., randomized, quasi-experiment with a no treatment comparison, and quasi-experimental with a treatment dropout comparison) were analyzed separately.

All studies were double-coded. Any differences were resolved by the two lead authors.

Statistical Procedures and Conventions

This systematic review used standard meta-analytic methods. More specifically, dichotomous program effects were encoded as odds-ratio type effect sizes (e.g., re-offend or not). Effects measured on continuous type measures (e.g., victim-reported abuse) were encoded as standardized mean difference type effect sizes (d). To ease presentation of the findings, all results were presented as standardized mean difference type effect sizes. This was accomplished by converting the odds-ratios into equivalent d -type effect sizes. This was done using a simple transformation that rescaled the logged odds-ratios (Haselblad & Hedges, 1995). This transformation does not affect the results in that significance levels, homogeneity statistics, etc. are unchanged. Effects representing unique constructs were analyzed separately (e.g., official report, victim report). The mean effect size across studies for any given construct was determined by weighting the inverse variance of the effect size, that is, using the inverse variance weight method. These analyses were performed with tools created by David B. Wilson that are publicly available (<http://mason.gmu.edu/~dwilsonb/ma.html>).

Treatment of Qualitative Research

This review did not synthesize the existing qualitative research in the area of domestic violence.

Results

Description of Studies

The above process identified 11,872 titles and abstracts (including duplicates). Fifty-seven studies were retrieved in their entirety for further scrutiny. Of these, a total of four experimental studies and six quasi-experimental studies were identified as meeting the eligibility criteria. The basic research design (i.e., randomized, quasi-experimental with a no treatment comparison group, quasi-experimental with treatment dropouts as the comparison group) and treatment type, number of treatment sessions and weeks, nature of the comparison group, and sample description are reported in Table 1.

All ten studies were conducted in North America. Nine of these studies were published in peer reviewed journals, although technical reports were also available for four studies (see reference list). When there was conflicting information between the two sources, data from the non-published technical report was used in the coding of the meta-analysis because that typically provided more detailed information.

All ten studies evaluated a psychoeducational or cognitive behavioral approach, or some mix of the two approaches targeted at the batterer and delivered in all-male group settings. One study (Dunford, 2000) also tested two additional intervention types: a cognitive behavioral group targeted at the male batterer but conducted in conjoint groups as well as a no-program but rigorously monitored intervention. In all but two of the studies (Chen et al., 1989; Dunford, 2000) it was noted that the program intervention was accompanied by probation, although in one of these studies (Chen et al., 1989) it seems likely that that was the case as well.

The treatment length ranged from a minimum of 8 two-hour sessions (Chen et al., 1989) to a maximum of 32 sessions over the course of a year (Dunford, 2000). Treatment length information was not provided by Syers and Edleson (1992). Many of the studies indicated the number of sessions and number of weeks but not the length of the treatment sessions.

The nature of the control group also varied from study to study. The Dunford study (2000) was the most unusual with the control group receiving no intervention whatsoever. Several studies (Feder & Forde, 2000; Gordon & Moriarty, 2003; Harrell, 1991; Palmer et al., 1992) had the control group receiving probation only. The Davis et al. study (2000) used a control group whose subjects received 40 hours of community services. The Gordon and Moriarty (2003) study included comparisons to both probation only and treatment no shows and drop-outs. Jones and Gondolf (2002) and Dutton (1986) also used treatment drop-out comparison group designs. Dutton (1986) included men who were rejected from treatment as well as the treatment no-shows and dropouts. Treatment no-shows and drop-outs represented 84% of the sample in Dutton's study and as such it is considered a treatment drop-out type study for the analyses below. Finally,

one study (Syers & Edleson, 1992) did not specify what the control group received, beyond not being mandated into counseling.

All but one of the ten studies used a general civilian population of batterers who were facing or had faced court prosecution for domestic violence. The one exception, Dunford (2000), used men living on a Navy base where an incident of domestic violence had been established and the man had been referred to the program. And all but one study (Jones & Gondolf, 2002) used a sample of men who were entirely court-mandated into the batterer program. The Jones and Gondolf (2002) study had a sample that was comprised of 79% court-mandated and 21% voluntary clients.

In five studies the generalizability of the sample to the general domestic violence offender population was questionable due to conditions used for inclusion into their sample. In one of the experimental studies (Palmer et al., 1992), inclusion criteria were suspected of being highly restrictive in that the resulting sample size was small despite the large jurisdiction from which it was pulled and the long time frame implemented for the study. A second experimental study (Davis et al., 2000) used highly restrictive criteria for inclusion in their sample. In that study, all individuals making up the courtroom workgroup, including the batterer, had to agree to this intervention (versus another non-jail alternative). This, as the researchers noted, led to a pool of more highly motivated offenders than is typically found in the generalized batterer population. In the Dunford study (2000), the men were all living on a naval base with their families and therefore may represent a group with higher stake in conformity than is true of other batterer samples. In one of the quasi-experimental studies (Syers & Edleson, 1992), only those men who could be followed six and twelve-month post-initial police visit were included in the study. This restriction makes it less likely that more marginal batterers would be included in their study. Another quasi-experimental study (Jones & Gondolf, 2002) excluded data from one of four sites because the men were deemed at higher risk for subsequent re-offending even though they had demonstrated a higher rate of completing treatment.

Meta-Analytic Synthesis of Studies

The effect sizes were analyzed separately by outcome type (official reports and victim reports) and by design type (experimental, quasi-experimental with a no-treatment comparison group, and quasi-experimental with treatment dropouts as the comparison group). Table 2 presents the random-effects mean effect size, 95% confidence interval, and homogeneity statistic (Q) for both outcome types and each design type. The results will be discussed separately for each outcome.

Official Reports: Official reports were either official complaints made to the police that may or may not have resulted in an arrest, or actual arrests for domestic violence. If multiple follow-up time points were available, the longest was selected. As can be seen from the Table 2, the mean effect size for the experimental (randomized) studies across these 7 comparisons was 0.26. This represents a finding of a moderate reduction in re-offending, with a 95% confidence interval of 0.03 to 0.50 ($z = 2.23$, $p = .03$). Figure 1 indicates a general pattern of positive effects on official reports of repeat victimization in these experimental studies. These estimates varied from a near zero effect (Davis et al., 8 week program) to large positive effects (Palmer et al, 10 week program; Davis et al., 26

week program). The mean represents a small positive reduction in repeat victimization. This effect roughly represents a reduction in recidivism from 20% to 13%. However, given the small number of studies (four), there is substantial uncertainty regarding the precision of this estimate.

There is additional doubt concerning what the results of one of these studies actually indicate. Specifically, though the Brooklyn Experiment was written as indicating modest support for court-mandated treatment's effectiveness (Davis et al, 2000; Taylor, Davis & Maxwell, 2001), the findings ran counter to expectations. As noted in their study, treatment completion was higher for the 8-week program than the 26-week program. Yet treatment effects were higher for the men assigned to the 26-week program with an effect size near zero for those assigned to the 8-week program. This differential effect suggests that something other than the batterer program accounted for the positive treatment effect. If the batterer program itself was effective, then the group receiving a higher dose (8-week program) should have had the better outcome. At the time, Feder and her colleagues speculated that these results were more consistent with a conclusion that supervision, and not treatment, resulted in the groups' differences in rates of reassault (Feder & Forde, 2000; Feder & Duggan, 2002). In fact, findings from the Brooklyn Experiment have now come to be viewed by the principal investigator as indicative that additional monitoring and not batterer programs were responsible for differences in recidivism between the three groups (Davis, personal communication). Despite this, the meta-analysis used the results as published by the Davis team (Davis et al., 2000). Thus, the strongest empirical evidence for the effectiveness of these programs comes from Palmer et al. (1992), a study with a very small sample size (30 men in the batterer program and 26 in the comparison condition). This small sample size leads to a very unstable estimate of the true treatment effect, as is evident in the rather large confidence interval.

We also noticed in coding the experimental studies that the offender population was restricted in some cases, that is, did not reflect the general domestic violence offender population in two studies—the Palmer and Davis studies (see Table 1). Analyzing the official report effect sizes by this distinction shows a lower non-significant overall mean effect size (0.12, with a 95% C.I. of -0.21 to 0.44) for the studies using a general domestic violence offender population and a higher mean effect size (0.39, with a 95% C.I. of 0.10 to 0.67) for the studies with a restricted sample. We are unsure what this finding suggests because the specific restrictions placed on one of these samples (Palmer et al, 1992) were not entirely clear. And, as discussed above, the pattern of results between the 8-week and 26-week programs, as well as Davis et al's own reinterpretation of their study's results (Davis et al., 2003) are not consistent with the hypothesis that batterer intervention programs are effective.

The quasi-experimental studies represented two fundamentally different design types: designs comparing offenders mandated to treatment to those not mandated and designs comparing treatment completers to treatment drop-outs, no-shows and/or rejects. Because the effect that each design is estimating is different, these two design types were analyzed separately. Table 2 indicates that the mean effect size across the former design (not mandated to treatment comparison) was -0.07, a small negative effect that is statistically not significantly different from zero. As indicated in Figure 2, these four credible quasi-experimental studies provide a mixed picture (also evidenced by the

significant homogeneity test, Q), with one study observing a moderate positive benefit, one a small positive benefit and two observing a negative effect of a court-mandated treatment relative to a non-mandated group. These estimates statistically adjust for baseline difference although it is unlikely that all of the important differences between the groups were taken into account. The composite or mean effect has a plausibility range extending from a small negative to a small positive effect.

The second quasi-experimental design type compared batterers who completed a court-mandated treatment program with those who were mandated and were either rejected from treatment, never showed or dropped out. The three studies with this design consistently found a positive and significant effect. That is, abusers mandated to a domestic violence treatment who complete their program re-offend at a substantially lower rate than offenders who were mandated to these programs who did not complete their treatment. Unfortunately, we cannot attribute this difference solely to the impact of treatment as treatment attendance is likely to be confounded with other important variables. That is, men who attend and complete their treatment may be more highly motivated to change or more fearful of further criminal justice involvement than are men who do not complete a treatment program that has been judicially mandated. Differences in rates of recidivism may be attributed, then, to differences existing in the groups prior to the intervention. In other words, the relationship that we think we are observing between treatment non-completion and recidivism may be spurious and due to another unobserved variable.

Victim Reported Outcomes: A concern with official measures is that they may not accurately reflect the amount and severity of on-going violence. Research consistently indicates that official reports capture only a small fraction of this abuse (Dutton, 1988; Straus, 1991; Tjaden & Thoennes, 2000). As such, the victim is viewed as the best source for information on the offender's continued abuse. Given that, we turn our attention to the seven estimates we have from these studies on the effect of these programs according to the victim's reports of abuse. Three of the four experimental studies measured the victim's reports of their partner's abusive behavior using either the standardized Conflict Tactics Scale (CTS) or the modified Conflict Tactics Scale (CTS2) (Straus, Hamby, Boney-McCoy, & Sugarman, 1996). One of the quasi-experimental studies also measured the victim's report of their partner's abusive behavior using a measure similar to the CTS. For purposes of analysis, we coded all reported subscales and averaged the multiple effect sizes within each treatment-comparison contrast, with the exception of Harrell (1991) where we selected the outcome based on the largest portion of the sample. Thus, the effect size used in Table 2 and Figure 4 represents the mean effect across subscales of the CTS/CTS2 for the comparison of interest. As shown in Table 2, the mean effect size for victim reports in studies using an experimental design was near zero and was not statistically significant. The effect size for quasi-experimental studies showed a small and negative effect for treatment though this finding also was not statistically significant. The distribution of effects is shown in Figure 4. Three of these effects are positive, four are negative, and none are statistically significant. Thus, the outcome measures based on the female intimate partner's report, and the more credible of the quasi-experimental studies, do not replicate the finding of a small but positive benefit of treatment found in the experimental studies using the official measures of re-offending.

Discussion

This systematic review was based on ten experimental and quasi-experimental studies. The experimental studies looked at the effect of mandating batterer intervention program relative to a no-treatment or routine-treatment approach for men facing or convicted of misdemeanor domestic violence charges. Two of the quasi-experimental studies compared men court-mandated to counseling with those not court-mandated (Syers & Edleson, 1992; Harrell, 1991), two compared men court-mandated who completed treatment to those mandated who did not complete treatment (Dutton, 1986; Snow and Gondolf, 2002) and one study (Gordon and Moriarty, 2003) included both comparisons. All of the evaluated programs used a psychoeducational, feminist oriented and/or cognitive behavioral approach.

The evidence from our meta-analysis is mixed. There is some support for the modest benefits of batterer programs from official reports in the experimental studies, but this effect is smaller (and non-significant) if we look only at studies using a general batterer population. Additionally, the effect is absent when victim reported measures are examined. The quasi-experimental studies using a no-treatment comparison also fail to find a positive treatment effect in terms of a reduction in violence when measured with official reports. Finally, quasi-experimental studies using men who were rejected from treatment or who rejected treatment were the only studies to consistently show a large, positive and significant effect on reducing re-offending.

It must be noted that we have serious concerns about these ten studies. The first of our four main concerns deals with the question of the generalizability of these findings to general convicted batterer populations. Second, we believe there is a potential bias inherent in using official records to measure continued abuse. Third, the victim reports suffer from low reporting rates in these studies raising concerns about the validity of the effect estimates from those studies. And finally, we question the validity of the quasi-experimental studies that compare treatment completers to rejects, no-shows, and dropouts. Each of these concerns is addressed below.

The issue of generalizability: We judged two studies (Davis et al., 2000; Palmer et al., 1992) as having samples that were restricted in a manner that reduced the generalizability of their findings to a general batterer population. Studies that did not have restrictions limiting who was included in the batterer program probably better represented the “typical” convicted batterer. Our analysis indicates that these latter studies had a lower overall mean effect size for official reports of domestic violence than the studies using a restricted sample. Importantly, the mean effect for the more representative studies was not statistically significant, raising the possibility that the overall positive finding of Figure 1 was in part a function of a restricted (possibly more motivated or perhaps “creamed”) sample of batterers. This may indicate that batterer intervention programs work for a selected (presumably more motivated) subset of offenders. The evidence on this issue is weak for two reasons: (1) we do not actually know the motivation levels of the men in the different studies, and (2) the Davis et al. study had inconsistent results across two similarly motivated groups receiving the same intervention, differing only in the number of weeks over which the program was spread. Thus, we believe that there is insufficient data for any strong conclusion on this issue.

The issue of heavy reliance on official reports: The heavy reliance on official measures in all of these studies is also highly problematic. Official measures are dependent on a victim's willingness to file a complaint or call the police. This raises the possibility that assignment to court-mandated treatment versus a no-treatment control group may differentially affect the victim's willingness to contact criminal justice officials when future abuse occurs (Cook and Campbell, 1979, refer to this as an instrumentation effect). A victim may not report her partner's abuse for a number of reasons. This includes the possibility that she might prefer to see her partner continue in treatment where she believes it will eventually lead to changes in his abusive behavior rather than take the risk of reporting his continued abuse and see him go to jail. Alternately, a victim may resent the criminal justice system's intrusion into her life in the form of mandating a treatment that she is then responsible for paying. Most programs require the abuser to pay for the treatment and by extension that means that it is the family that pays for the treatment (Zorza, 2003). If the treatment is viewed by a victim as ineffective, it may make her critical and suspicious of the system and less likely to cooperate in the case of reporting future incidences of abuse. We have no empirical evidence that this occurs, but the dependence of official reports on the behavior of the victim allows for the plausibility that the different rates noted between batterers in the treatment and comparison conditions may reflect a measurement artifact and not a genuine treatment effect. This possibility is strengthened by the different findings obtained in these studies depending upon whether official reports or victim reports are used as the outcome measure.

The issue of the low victim reporting rate: The high rate of victim attrition in many of these studies is another concern. The victim is usually viewed as the best source for information on the offender's continued abuse. Victim reports of abuse via standardized measures such as the Conflict Tactics Scale are less likely to be affected by the issues raised regarding official reports of continued abuse, provided that the victim is convinced of the confidential nature of her responses. Unfortunately, the percentage of victims responding to follow-up surveys in these studies is low, seriously undermining their utility in establishing the effectiveness (or ineffectiveness) of these programs.

The attrition for victim report for the effect sizes shown in Figure 4 was roughly 30% for the Dunford (2000b), roughly 50% for the Davis et al. (2000), roughly 80% for the Feder and Dugan (2002) and 59% for Harrell (1991). High attrition raises the possibility that the victims lost to follow-up in the treatment group may differ in meaningful ways from those lost to follow-up in the control group. Thus, the absence of an effect for the victim report measures may reflect that the programs are truly ineffective or, alternately, that there is a positive or negative effect that is masked by differential attrition.

The problem of high rates of victim attrition becomes critical in light of research indicating that certain victims of domestic violence are more likely to be lost in the research follow-up than are others. This research strongly suggests that women victims of domestic violence who are more difficult to retain in follow-up research are both more marginal and more likely to be more frequently and severely abused (Sullivan, Rumpitz, Campbell, Eby and Davidson, 1996). There is also research that indicates that men who are more marginal are both less likely to obey a court-mandate to treatment and more likely to continue to abuse their partners (Feder & Dugan, 2002). If we can assume that more marginal women are more likely to be partnered with more marginal men, then the need for maintaining contact with a high percentage of victims when assessing the

effectiveness of these spouse abuse abatement programs becomes even more apparent. This may be important to the extent that some research has indicated that factors associated with the abuser's stake in conformity is associated with the likelihood that an intervention will be successful in reducing subsequent violence (Berk, Campbell, Klap and Western, 1992; Sherman, 1992). At best, this attrition reduces the generalizability of the findings from victim reported outcomes to a subset of the domestic violence offender population. At worst, there may be differential loss of these marginal women from the treatment and control groups, producing bias in the findings.

The issue of the validity of using treatment drop-outs as a comparison: Finally, we note the difficulty with using treatment dropouts as a control group even once statistical controls have been introduced. Two specific problems occur with this type of study design, one with the construct of what is being evaluated or tested and the other with the adequacy of the statistical models in adjusting for initial group differences. First, these studies are trying to estimate the affect of full participation in the batterer intervention program above and beyond the court mandate. In other words, they look to answer the question, "Among men who are court-mandated to batterer intervention, do those who choose to attend and complete this program do better than those who do not?" Although this may be of interest to program providers and developers, it does not address the broader issue of the likely reduction in domestic violence as a function of a policy to mandate such treatment. That is, "whether court-mandated batterer intervention programs reduce offenders' likelihood of re-offending." Addressing the latter question is critical to knowing whether court-mandated domestic violence interventions are beneficial to society.

Second, that these studies produce treatment effect estimates that are large given the population and nature of the problem clearly establishes that men who complete these programs recidivate at a lower rate than men who do not. The question is what to make of this empirical finding. The statistical models employed by these studies attempt to adjust for selection differences between the groups of men. To produce unbiased estimates, however, these models need to fully account for the selection process, that is, the reasons why some men attended treatment and others did not. We do not believe that these equations adequately model the selection process. Potentially important variables, such as motivation for treatment, were not included. The positive treatment effect estimate may indicate that the treatment is effective for motivated offenders though we cannot conclude this since we do not have comparisons in any of these quasi-experimental studies with motivated offenders who were not mandated and did not receive treatment. Equally plausible, these findings may simply reflect that the subset of offenders who will complete mandated treatments are less likely to re-offend, with or without the treatment (i.e., these programs may have "creamed" those offenders who are least likely to re-offend regardless of what action is taken).

Our findings are somewhat different from those of Babcock, et al (2004). They concluded, based on their meta-analysis, that these programs have a small but positive effect on abusive behavior. There are several differences between the methods employed in our respective meta-analysis that may account for the differing conclusions. Primarily, Babcock et al. did not separately analyze studies using treatment drop-out designs from

other quasi-experimental designs, potentially upwardly biasing the mean effect size for these studies (as discussed in detail above). If one looks only at experimental studies, results from both meta-analyses are fairly consistent. Babcock et al. reported an effect size of 0.12 when using official reports (fixed effects 95% C.I. of 0.02 to 0.22). This is somewhat smaller than our overall mean effect for official reports based on experimental studies but consistent with our estimate from those studies with a representative population. Similarly, Babcock and Steiner indicate a treatment effect of 0.09 (fixed effects 95% C.I. of -0.02 to 0.21) for victim reported outcomes, slightly higher than our estimate (0.01) but neither estimate is statistically significant.

Reviewers' Conclusions

The findings from this meta-analysis combined with the caveats above raise questions as to the value of these programs. While additional research is needed, the meta-analysis does not offer strong support that court-mandating treatment to misdemeanor domestic violence offenders reduces the likelihood of further reassault.

Implications for Practice

Intervening in the lives of others is a risky business, particularly when the individuals participating in the social intervention are mandated by a court of law to do so. As such, it is incumbent upon us to ensure that we are not inadvertently making things worse for those we are seeking to help. At this point the existing evidence cannot ensure that these programs are, in fact, helpful and not harmful.

There is no doubt that, "There is a tremendous sense of urgency and alarm in the treatment of domestic violence—and rightly so. After all, protecting the physical and emotional safety of women and their children is the first priority. Consequently, clinicians feel a primary obligation to 'do something' immediately and decisively to halt and prevent violence" (Jennings, 1987, p. 204). But as the above review has indicated, doing something may not help. As McCord so wisely noted, "Unless social programs are evaluated for potential harm as well as benefit, safety as well as efficacy, the choice of which social programs to use will remain a dangerous guess" (McCord, 2003, p. 16). It is clear that we need to be guided by rigorous research in helping us set our course.

While better research is needed to determine the effectiveness of court-mandated batterer intervention programs, the results from the meta-analysis do not provide confidence that these programs will be found to be effective. Therefore, it would prove beneficial for the criminal justice system to begin looking at other types of interventions for addressing the problem of domestic violence. But these interventions must be tied to rigorous evaluations to determine their full impact. In other words, we recommend the use of pilot studies joined to an experimental design, as was suggested almost twenty years ago by Berk and his colleagues (Berk, Boruch, Chambers, Rossi & Witte, 1985), as the preferred path for finding effective programs that can meet the challenge that intimate partner violence presents. Such a course would be especially prudent in these times of limited resources. More than that, victims and taxpayers are deserving of such evidence-based decision-making.

Unfortunately, what we are suggesting is not possible in many jurisdictions today in that their statutes require that, upon conviction for domestic violence, individuals must be

mandated into a batterer intervention program, not atypically based upon the Duluth Model (Babcock & Taillade, 2000). The end result is that judges, prosecutors and probation officers continue to send batterers to these treatment programs even as they have grave doubts about their effectiveness. And alternate programs cannot be implemented and tested even as evidence builds indicating that batterer intervention programs, at least as designed and implemented today, may not be effective.

Implications for Research

The research implication growing out of this synthesis is that additional experiments need to be conducted to more clearly decipher the effectiveness of court-mandated batterer intervention programs. If we are to test the ability of courts to mandate change, these future experiments must ensure samples of batterers that are representative of the larger convicted batterer population rather than a smaller subset of selected batterers. Additionally, these studies must attend to the importance of maintaining high victim retention so as to better ascertain any positive or negative effects from this mandated intervention. Finally, additional research is needed to better understand the validity and reliability of official report and victim report measures used in these studies and how they might be affected by treatment assignment.

Plans for Updating the Review

The two lead authors will be responsible for updating the review every three years if needed.

Acknowledgements

We are grateful for the Smith-Richardson Foundation for their support of this project.

Potential Conflict of Interest

The first listed author (Lynette Feder) was the primary investigator of an experiment assessing the effectiveness of a court-mandated counseling program conducted in Broward County, Florida. To best counter the potential conflict of interest, the review was made as transparent as possible and included a collaborator who had not been involved in any of the prior research reviewed here.

References

Studies Included in this Review

- Chen, H, Bersani, C., Myers, S., & Denton, R. (1989). Evaluating the effectiveness of a court sponsored treatment program. *Journal of Family Violence*, 4, 309-322.
- Davis, R. C., Taylor, B. G., & Maxwell, C. D. (2000). *Does batterer treatment reduce violence? A randomized experiment in Brooklyn*. Washington, DC: National Institute of Justice.
- Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68, 468-476.

- Dutton, D. (1986). The outcome of court-mandated treatment for wife assault: A quasi-experimental evaluation. *Violence and Victims*, 1(3), 163-175.
- Feder, L. & Forde, D. (2000). *A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward Experiment* (Final report, Grant NIJ-96-WT-NX-0008). Washington, DC: National Institute of Justice.
- Feder, L. & Dugan, L. (2002). A test of the efficacy of court mandated counseling for domestic violence offenders: The Broward Experiment. *Justice Quarterly*, 19(2), 343-375.
- Gordon, J. A. & Moriarty, L. J. (2003). The effects of domestic violence batterer treatment on domestic violence batterer treatment on domestic violence recidivism. *Criminal Justice and Behavior*, 30(1), 118-134.
- Harrell, A. (1991). *Evaluation of court-ordered treatment for domestic violence offenders* (Final report). Washington, D.C.: National Institute of Justice.
- Jones, A. S. & Gondolf, E. W. (2002). Assessing the effect of batterer program completion on reassault: An instrumental variables analysis. *Journal of Quantitative Criminology*, 18(1), 71-98.
- Palmer, S., Brown, R., & Barrera, M. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry*, 62(2), 276-283.
- Syers, M. & Edleson, J. (1992). The combined effects of coordinated criminal justice intervention in woman abuse. *Journal of Interpersonal Violence*, 7, 490-502.

Studies Excluded from this Review

- Alaska Judicial Council (1999). *Evaluation of Pilot Program for Misdemeanor Domestic Violence Offenders* (Final Report). [city, state:] Author.
- American Medical Association (1995). *A coordinated approach to reducing family violence: conference highlights*. [city, state: publisher].
- Aubertin, N., & Laporte, P. (1999). Contrecoups: A program of therapy for spousal and family violence. *Forum on Corrections Research*, 11(1), 3-5.
- Babcock, J., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology*, 13(1), 46-59.
- Barrera, M. E., Palmer, S. E., Brown, R. A., & Kalher, S. (1994). Characteristics of court-involved and non-court-involved men who abuse their wives. *Journal of Family Violence*, 9(4), 333-345.
- Dobash, R., Dobash, R., Cavanagh, K., & Lewis R. (1996). *Re-education Programmes for violent men: an evaluation*. Edinburgh: HMSO

- Dobash, R., Dobash, R., Cavanagh, K., et al. (1996). Research evaluation of programmes for violent men. In Dobash R., et al. (Eds.), *Changing violent men*. Thousand Oaks, CA: Sage Publications
- Dobash, R. P., Dobash, R. E., Cavanaugh, K., & Lewis, R. (1999). A Research Evaluation of British Programmes for Violent Men. *Journal of Social Policy*, 28(2), 205-233.
- Dobash, R. E., & Dobash, R. P. (2000). Evaluating criminal justice interventions for domestic violence. *Crime and Delinquency*, 46(2), 252-270.
- Dobash, R. P., Dobash, R. E., Cavanagh, K., & Lewis, R.. (2000). Confronting violent men. In J. Hanmer, & C. Itzin (Eds.), *Home truths about domestic violence: Feminist influences on policy and practice a reader* (pp. 289-309). London, Routledge.
- Dutton, D., Bodnarchuk, M., Kropp, R., Hart, S., & Ogloff, J. (1997). Wife assault treatment and criminal recidivism: An 11 year follow-up. *International Journal of Offender Therapy and Comparative Criminology*, 41(1), 9-23.
- Edleson, J. L., & Grusznski, R. J. (1988). Treating men who batter: Four years of outcome data from the Domestic Abuse Project. *Journal of Social Science Research*, 12(1-2), 3-22.
- Flournoy, P. (1993). A comparison of groups for men who batter. (Doctoral dissertation, Washington State University) *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 61(9-B), 4989.
- Ford, D. A. (1991). Preventing and provoking wife battery through criminal sanctioning: A look at the risks. In D. D. Knudsen, & J. L. Miller (Eds.), *Abused and battered: Social and legal responses of family violence. Social Institutions and social change* (pp. 191-209). New York, NY: Aldine de Gruyter
- Ford, D., & Regoli, M. (1992). The preventative impacts of policies for prosecuting wife batterers. In E. S. Buzawa, & C. G. Buzawa (Eds.), *Domestic violence: The changing CJ response* (pp. 181-208). Dover, MA: Auburn House
- Ford, D., & Regoli, M. (1993). *The Indianapolis domestic violence prosecution experiment* (Final Report). Rockville, MD: National Institute of Mental Health. (NIMH #15161-13)
- Ford, D., & Regoli, M. (1993). The Criminal Prosecution of Wife Assaulters: Process, Problems & Effects. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation*. Newbury Park: Sage Publications
- Gamache, D., Edleson, J., & Schock, M. (1988). Coordinated police, judicial, and social service response to woman battering: A multiple-baseline evaluation across three communities. In G. Hotaling, & D. Finkelhor, et al. (Eds.), *Coping with family violence: Research and policy perspectives* (pp. 193-209). Newbury Park: Sage Publications

- Goldkamp, J., Weiland, D., Collins, M. et al. (1996). *The role of drug and alcohol abuse in domestic violence and its treatment: Dade County's Domestic Violence Court Experiment* (Final Report). Washington, DC: National Institute of Justice. (NCJRS #163410)
- Gondolf, E. W., & Jones, A. S. (2001) The program effect of batterer programs in three cities. *Violence and Victims*, 16(6), 693-704.
- Gondolf, E. W. (2000). A 30-month follow-up of court referred batterers in four cities. *International Journal of Offender Therapy & Comparative Criminology*, 44 (1), 111-128.
- Gondolf, E. W. (1999). A comparison of four batterer intervention systems: Do court referral, program length, and services matter? *Journal of Interpersonal Violence*. 14(1) Jan, 41-61.
- Gondolf, E. W. (1988). How some men stop their abuse: An explanatory programs evaluation. In G. T. Hotaling, & D. Finkelhor, et al. (Eds.), *Coping with family violence: Research and policy perspectives* (pp. 129-144). Newbury Park: Sage Publications
- Gondolf, E. W. (1997). Patterns of Reassault in batterer programs. *Violence and Victims*, 12(4), 373-387.
- Hamberger, K., Hastings, J. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3(2), 121-1130.
- Hamm, M., & Kite, J. (1991). The role of offender rehabilitation in family violence policy: the Batterers Anonymous Experiment. *Criminal Justice Review*, 16(2), 227-248.
- Heckert, D. A., & Gondolf, E. W. (2000). The effect of perceptions of sanctions on batterer program outcomes. *Journal of Research in Crime and Delinquency*, 37(4), 369-391.
- Jolin, A., Feyerherm, W., Fountain, R., & Friedman, S. (1998). *Beyond arrest: The Portland, Oregon Domestic Violence Experiment* (Final Report). Washington, DC: U.S. Department of Justice. (NCJRS #179968)
- Kistenmacher, B. R. (2001). Motivational Interviewing as a mechanism for change in men who batter: a randomized controlled trial. (Doctoral dissertation, University of Oregon) *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 61(9-B). 4989.
- Krmpotich, S., & Eckberg, D. (2000). *Domestic assault program evaluation: Final (2-year) results*. Minneapolis, Hennepin County Department of Community Corrections.
- Ley, D.J. (2001). Effectiveness of a court-ordered domestic violence treatment program: a clinical utility study. (Doctoral dissertation, University of New Mexico). *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 62(4-B). 2056.

- Morrell, T., Elliott, J., Murphy, C., & Taft, C. (2003). Cognitive behavioral and supportive group treatment for partner-violent men. *Behavior Therapy*, 34, 77-95.
- Murphy, C. M., Musser, P. H., & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. *Journal of Family Violence*, 13(3), 263-284.
- National Institute of Justice (1998). *Legal interventions in family violence: research findings and policy implications* Washington, U.S. Department of Justice.
- Healey, K. and Smith, C. (1998). Batterer programs: what criminal justice agencies need to know. Washington: U.S. Department of Justice .
- Newell, R. G. (1994). The effectiveness of court-mandated counseling for domestic violence: An outcome study. (Doctoral dissertation, University of Toledo. *Dissertation Abstracts International: Section A: the Sciences & Engineering*, 55(05), 1193.
- Pellegrini, K. L. (1999). Analysis of a violence intervention program: Population, treatment compliance, and recidivism. (Doctoral dissertation, George Fox University) *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 60(10-B), 5231.
- Petrik, N., Gildersleeve-High, L., McEllistrem, J. (1994). The reduction of male abusiveness as a result of treatment: reality or myth? *Journal of Family Violence*, 9, 307-316.
- Petrucci, C. J. (2002) A qualitative & quantitative analysis of a specialized DV court that utilizes therapeutic jurisprudence. (Doctoral dissertation, University of California, Los Angeles.
- Taft, C., Murphy, C., Elliot, J. & Morrell, T. (2001). Attendance-enhancing procedures in group counseling for domestic abuse. *Journal of Consulting Psychology*, 48(1), 51-60.
- Taylor, B., Davis, R., Maxwell, C. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly*, 18(1), 171-201.
- Tolman, R. M., Bhosley, G. (1991). The outcome of participation in a shelter- sponsored program for men who batter. In D. D. Knudsen, & J. L. Miller (Eds.), *Abused and Battered: Social and legal responses of family violence. Social Institutions and social change* (pp. 191-209). New York: Aldene de Gruyter.
- Tutty, L. M., Bidgood, B. A., Rothery, M. A., & Bidgood, P. (2001). An evaluation of men's batterer treatment groups. *Research on Social Work Practice*, 11(6), 645-670.
- Waldo, M. (1988). Relationship enhancement counseling groups for wife abusers. *Journal of Mental Health Counseling*, 10(1), 37-45.

Other References Cited in the Text

- Adams, D. & McCormick, A. (1982). Men unlearning violence: A group approach based on the collective model. In M. Roy (Ed.), *The abusive partner: An analysis of domestic battering* (pp. 170-197). New York: Van Nostrand Reinhold.
- Babcock, J. C. & Taillade, J. (2000). Evaluating interventions for men who batter. In J. Vincent & E. Jouriles (Eds.), *Domestic violence: Guidelines for research-informed practice* (pp. 37-77). Philadelphia: Jessica Kingsley.
- Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, 23(8), 1023-1053.
- Berk, R., Campbell, A., Klap, R., & Western, B. (1992). The deterrent effect of arrest in incidents of domestic violence: A Bayesian analysis of four field experiments. *American Sociological Review*, 57(5), 698-708.
- Berk, R., Boruch, T., Chambers, F., Rossi, P., & Witte, S. (1985). Social policy experimentation: A position paper. *Evaluation Review*, 9(4), 387-429.
- Brisson, N. (1981). Battering husbands: A survey of abusive men. *Victimology*, 6, 338-344.
- Chalk, R. & King, P. (1998). *Violence in families: Assessing prevention and treatment programs*. Washington, DC: National Academy Press.
- Davis, R. & Taylor, B. (1999). Does batterer treatment reduce violence? *Women & Criminal Justice*, 10, 69-93.
- Davis, R., Maxwell, C., & Taylor, B. (2003). The Brooklyn Experiment. In S. Jackson, L. Feder, D. Forde, R. Davis, C. Maxwell & B. Taylor (Eds.), *Batterer intervention programs: Where do we go from here?* (pp. 15-21). Washington, DC: Department of Justice.
- Dutton, D. (1984). Interventions into the problem of wife assault: Therapeutic, policy and research implications. *Canadian Journal of Behavioral Science*, 16(4), 281-297.
- Dutton, D. (1988). Research advances in the study of wife assault: Etiology and prevention. *Law and Mental Health*, 4, 161-220.
- Dutton, D. & McGregor, B. (1991). The symbiosis of arrest and treatment for wife assault: The case for combined intervention. In M. Steinman (Ed.), *Woman battering: Policy responses* (pp. 131-154). Cincinnati, Ohio: Anderson Publishing Company.
- Eisikovits, Z. & Edleson, J. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review*, 63, 384-414.
- Edleson, J. & Brygger, M. (1995). Gender differences in reporting of battering incidences. In S. Stith & M. Straus (Eds.), *Understanding partner violence: Prevalence, causes, consequences and solutions*, (pp. 45-50). Minneapolis, MN: National Council of Family Relations.

- Farley, D. & Magill, J. (1988). An evaluation of a group program for men who batter. *Social Work With Groups*, 11(3), 53-65.
- Feazell, C., Mayers, R., & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations*, 33, 217-223.
- Feder, L. (1997). Domestic violence and police response in a pro-arrest jurisdiction. *Women and Criminal Justice*, 8(4), 79-98.
- Ford, D. & Regoli, M. J. (1993). The criminal prosecution of wife assaulters. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 127 -164). Newbury Park, CA: Sage.
- Gondolf, E. (1987). Seeing through smoke and mirrors: A guide to batterer program evaluations. *Response*, 10, 16-19.
- Gondolf, E. (1987). Evaluating programs for men who batter: Problems and prospects. *Journal of Family Violence*, 2(1), 95-108.
- Gondolf, E. (1998). Do batterer programs work? A 15 month follow-up of multi-site evaluation. *Domestic Violence Report*, 3(5), 65-80.
- Goolkasian, G. (1986). *Confronting domestic violence: The role of criminal court judges*. Washington, DC: National Institute of Justice.
- Hamberger, L. K. & Hastings, J. (1989). Counseling male spouse abusers: Characteristics of treatment completers and dropouts. *Violence and Victims*, 4(1), 275-286.
- Hamberger, L. K. & Hastings, J. (1993). Court-mandated treatment of men who assault their partner. In Z. Hilton (Ed.), *Legal Responses to Wife Assault: Current Trends and Evaluation* (pp. 188-229). Newbury Park, CA: Sage.
- Hasselblad, V. & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*, 117, 167-178.
- Healey, K., Smith, C., & O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies*. Washington, DC: Department of Justice.
- Healey, K. & Smith, C. (1998). *Batterer programs: What criminal justice agencies need to know*. Washington, DC: National Institute of Justice.
- Hilberman, E. (1980). Overview: The "wife-beater's wife" reconsidered. *American Journal of Psychiatry*, 137(11), 1336-1347.
- Hirschel, J. D. & Hutchinson, I. (1992). Female spouse abuse and the police response: The Charlotte, North Carolina Experiment. *Journal of Criminal Law and Criminology*, 83(1), 73-119.
- Hotaling, G. & Sugarman, D. (1986). An analysis of risk markers in husband to wife violence: The current state of knowledge. *Violence and Victims*, 1(2), 101-124.
- Jennings, J. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence*, 2(3), 193-213.

- Johnson, J. & Kanzler, D. (1993). Treating domestic violence: Evaluating the effectiveness of a domestic violence diversion program. *Studies in Symbolic Interaction, 15*, 271-289.
- Langan, P. & Innes, C. (1986). *Preventing domestic violence against women*. Washington, DC: National Institute of Justice.
- McCord, J. (2003) Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. *Annals of the American Academy of Political and Social Science, 587*, 16-30.
- Miller, T., Cohen, M., & Wiersema, B. (1996). *Victim costs and consequences: A new look*. Washington, DC: National Institute of Justice.
- Pence, E. (1983). The Duluth Domestic Abuse Intervention Project. *Hamline Law Review, 6*, 247-275.
- Pirog-Good, M. & Stets-Kealey, J. (1985). Male batterers and battering prevention programs: A national survey. *Response, 8*, 8-12.
- Rennison, C. R. & Welchans, S. (2000). *Intimate partner violence*. Washington, DC: National Institute of Justice.
- Roberts, A. (1982). A national survey of services for batterers. In Maria Roy (Ed.), *The Abusive Partner: An Analysis of Domestic Battering* (pp. 230-243). New York: Van Nostrand Reinhold.
- Rosenfeld, B. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review, 12*, 205-226.
- Saunders, D. (1996). Interventions for men who batter: Do we know what works? *In Session: Psychotherapy in Practice, 2*(3), 81-93.
- Sherman, L. (1992). The influence of criminology on criminal law: Evaluating arrests for misdemeanor domestic violence. *Journal of Criminal Law and Criminology, 83*, 1-45.
- Snyder, D. & Scheer, N. (1981). Predicting disposition following brief residence at a shelter for battered women. *American Journal of Community Psychology, 9*, 559-566.
- Sonkin, D. J. (1988). The male batterer: Clinical and research issues. *Violence and Victims, 3*(1), 65-79.
- Straus, M. (1991). Conceptualization and measurement of battering: Implications for public policy. In M. Steinman (Ed.), *Woman Battering: Policy Responses* (pp. 19-47). Cincinnati, Ohio: Anderson.
- Straus, M., Hamby, S., Boney-McCoy, S., & Sugarman, D. (1996). The revised Conflict Tactics Scale (CTS2): Development and preliminary psychometric data. *Journal of Family Issues, 17*(3), 283-316.
- Sullivan, C., Rumpitz, M., Campbell, R., Eby, K., & Davidson, W. (1996). Retaining participants in longitudinal community research: A comprehensive protocol. *Journal of Applied Behavioral Science, 32*(3), 262-276.

- Taylor, B., Davis, R., & Maxwell, C. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly*, 18(1), 171-201.
- Tjaden, P. & Thoennes, N. (2000). Prevalence and consequences of male-to-female and female-to-male intimate partner violence as measured by the National Violence Against Women Survey. *Violence Against Women*, 6(2), 142-161.
- Tolman, R. & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence*, 5, 87-118.
- Tolman, R. & Edleson, J. (1995). Intervention for men who batter: A review of research. In S. Stith & M. Straus (Eds.), *Understanding partner violence: Prevalence, causes, consequences and solutions* (pp. 262-273). Minneapolis, MN: National Council on Family Relations.
- Weisburd, D., Lum, C., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *Annals of the American Academy of Political and Social Science*, 578, 50-70.
- Widom, C. S. (1992). *The cycle of violence*. Washington, DC: US Department of Justice.
- Zorza, J. (2003). New research: Broward County Experiment shows no benefit from batterer intervention programs. *Domestic Violence Report*, 8, 23-25.

Tables

Table 1

Description of Studies by Author and Design Type

Author by Design Type	Treatment Type	Treatment Sessions/Weeks	Comparison Type	Sample Type
Randomized				
Davis et al.—8 week program	Psychoeducational	16/8	Probation and 40 hours community service	Convicted batterers—Judge, prosecutor & defense must agree to treatment
Davis et al.—26 week program	Psychoeducational	26/26	Same as above	Same as above
Dunford—Men's group	Cognitive-behavioral	32/52	No treatment	Navy sample, incident of domestic violence established, referred to program
Dunford—Conjoint	Cognitive-behavioral	32/52	No treatment	Same as above
Dunford—Rigorous monitoring		12/52	No treatment	Same as above
Feder & Forde	Cognitive-behavioral/ Psychoeducational	26/26	Probation	All convicted batterers
Palmer et al.	Psychoeducational	10/10	Probation	Convicted batterers—Unclear how sample drawn

Quasi-Experimental—No Treatment Comparison

Chen et al.	Cognitive-behavioral/ Psychoeducational	8 sessions	Non-referred convicted batterers	Convicted batterers referred to treatment program—Unclear how sample drawn
Gordon & Moriarty—Mandated vs. not	Psychoeducational	22/22	Probation	All convicted batterers
Harrell	Cognitive-behavioral	10/10	Probation	All batterers convicted or given prosecution deferred
Syers & Edleson	Psychoeducational		Batterers not mandated to counseling	All batterers having police contact who could be followed for 12 months

Quasi-Experimental—Dropouts as Comparison

Dutton	Cognitive-behavioral	16/16	Treatment dropouts, no-shows and rejects	Convicted batterers—Unclear how sample drawn
Jones & Gordon	Cognitive-behavioral	20/20	Treatment dropouts	Batterers in 4 treatment programs 79% court-mandated/21% voluntary referrals
Gordon & Moriarty—Completers vs. Dropouts	Psychoeducational	22/22	Treatment dropouts	All convicted batterers

Notes: Distinct treatment-comparison contrasts within an individual study are listed separately.

Table 2

Random effects mean effect size (d) and related statistics for official and victim reported measures of domestic violence by design type

	95% C. I.					
Outcome by Design Type	Mean <i>d</i>	Lower <i>d</i>	Upper <i>d</i>	<i>k</i> ^a	Q	τ ²
Official Measures						
Experiments (Randomized)	0.26*	0.03	0.50	7	8.19	.0256
Quasi-Experiments (Nonrandomized)						
No Treatment Comparison Group	-0.07	-0.45	0.31	4	12.00*	.1091
Treatment Dropouts as Comparison ^b	0.97*	0.12	1.82	3	12.00*	.4595
Victim Report Measures						
Experiments (Randomized)	0.01	-0.11	0.13	6	1.84	.0000
Quasi-Experiments (No Treatment Comp.) ^c	-0.11	-0.50	0.27	1		
Total	-0.00	-0.12	0.11	7	2.18	.0000

* $p \leq .05$

a. Number of effect sizes.

b. Fixed effects mean effect size was lower (mean $d = 0.49$, 95% C.I. of 0.27 to 0.71). Although substantially lower in value, this still represents a large effect in this context and leads to the same substantive conclusions.

c. Fixed effect

Figures

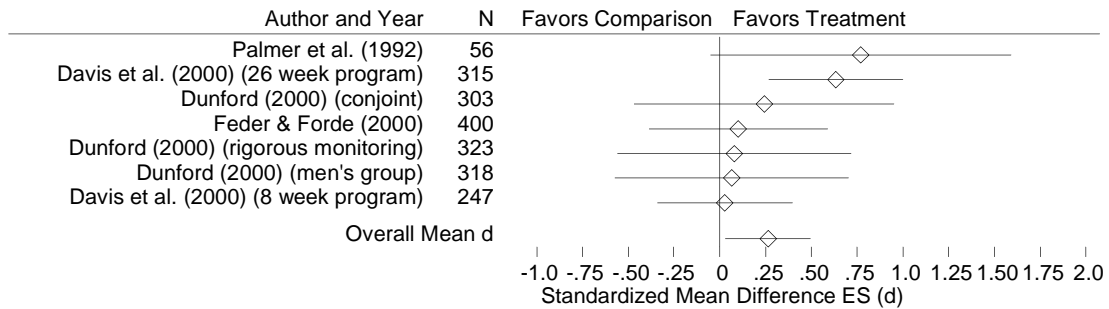


Figure 1: Effect Size (d) and 95% Confidence Interval for Official Measure from Experimental (Random) Studies

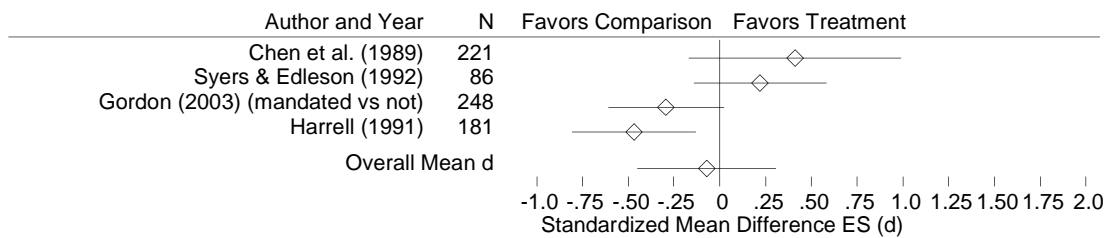


Figure 2: Effect Size (d) and 95% Confidence Interval for Official Measures from Quasi-Experimental (Nonrandomized) Studies with a No Treatment Comparison Group

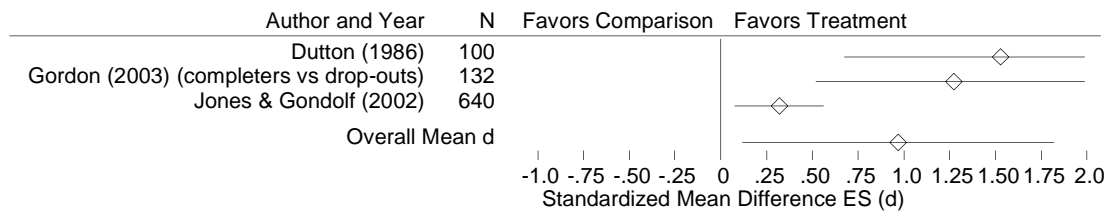


Figure 3: Effect Size (d) and 95% Confidence Interval for Official Measures from Quasi-Experimental (Nonrandomized) Studies with Treatment Dropouts as Comparison

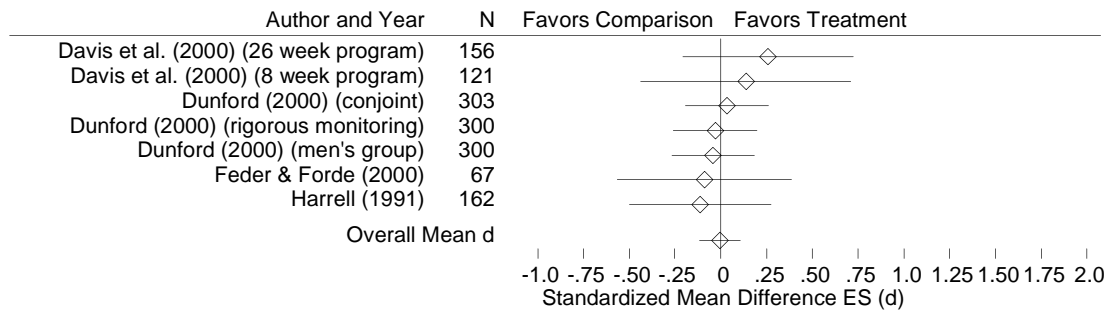


Figure 4: Effect Size (d) and 95% Confidence Interval for Victim Reported Measures from Experimental and Quasi-Experimental Studies with a No Treatment Comparison Group

Appendix One: Databases and Websites To Be Searched

Published

PsycINFO (Using OVID)

PsycINFO covers the professional and academic literature in psychology and related disciplines including medicine, psychiatry, nursing, sociology, education, pharmacology, physiology, and linguistics.

Coverage: 1889 - present

ERIC (Using OVID)

ERIC indexes journal articles and documents in the field of education and includes information from RIE (Resources in Education) and CIJE (Current Index to Journals in Education).

Coverage: 1966 – present

MEDLINE (Using OVID)

MEDLINE covers the international literature on biomedicine, including the allied health fields and the biological and physical sciences, humanities, and information science as they relate to medicine and health care.

Coverage: 1965 – present

Sociological Abstracts (Using OVID)

Sociological Abstracts covers the academic and professional literature in sociology and related disciplines in the social and behavioral sciences. The database draws information from an international selection of over 2,600 journals and other serials publications, plus conference papers, books, and dissertations. Citations to journal articles published after 1974 include abstracts.

Coverage 1963 - present

Social Science Citation Index

Provides citations from ~520 English-language periodicals in the areas of anthropology, criminology, economics, law, geography, policy studies, psychology, sociology, social work, and urban studies. Journals covered include the major periodicals in these disciplines.

Coverage: 1983 - present

Lexis Nexis Academic Universe

Provides a variety of full text sources including: regional, national, and international newspapers; magazines; wire services; business publications (trade journals, corporate annual reports, tax sources); legal resources (law reviews, court cases, briefs, federal and state codes); government documents; medical information (medical journals); and reference sources (directories, biographical information).

Social Work Abstracts

Social Work Abstracts, published by the National Association of Social Workers, provides access to the Social Work Abstracts and the Register of Clinical Social Workers. The database provides indexing and abstracting drawn from approximately 450 journals in all areas of social work, including theory and practice, areas of service, social issues, and social problems. The Register of Clinical Social Workers is a directory of clinical social workers in the United States.

Coverage: 1977 - present

Criminal Justice Abstracts

Criminal Justice Abstracts provides citations and abstracts of international journals, books, reports, dissertations and unpublished papers on criminology and related disciplines. Prepared in co-operation with the Criminal Justice Collection of Rutgers University Library, Criminal Justice Abstracts covers crime trends, crime prevention and deterrence, juvenile delinquency, juvenile justice, police, courts, punishment and sentencing.

Coverage: 1968 – present

Non-Published

Dissertation Abstracts International

"The database includes citations for materials ranging from the first U.S. dissertation, accepted in 1861, to those accepted as recently as last semester; those published from 1980 forward also include 350-word abstracts, written by the author. Citations for master's theses from 1988 forward include 150-word abstracts. The database represents the work of authors from over 1,000 North American graduate schools and European universities."

Coverage: 1861 –present

Governmental

GPO Monthly Catalog (MOCAT)

The Monthly Catalog of United States Government Publications (MOCAT) contains most of the information available through the Federal Depository Library Program and is available for online searching as part of the GPO Access Federal Locator services.

Coverage: 1994- updated daily

National Criminal Justice Research Service

The National Criminal Justice Reference Service Abstracts Database is produced by NCJRS, a service of the National Institute of Justice, with the Office of Juvenile Justice and Delinquency Prevention, Office for Victims of Crime, Bureau of Justice Statistics, and Bureau of Justice Assistance, all part of the Office of Justice Programs, U.S. Department of Justice, and the Office of National Drug Control Policy.

The database contains summaries of more than 150,000 publications on criminal justice, including Federal, state, and local government reports, books, research reports, journal articles, and unpublished research. Subject areas include corrections, courts, drugs and crime, law enforcement, juvenile justice, crime statistics, and victims of crime.

Coverage: Early 1970's - the present

UK National Health Service NRR (National Research Register)

<http://www.update-software.com/nrr/CLIBINET.EXE?A=1&U=1001&P=10001>

Existing Bibliographies

Babcock & La Taillade 1999

Babcock, J.C. & La Taillade, J.J. Research on treatment of men who batter: A synthesis of the outcome of literature and recommendations. Domestic Violence: Guidelines for Research-Informed Practice. Philadelphia, PA: Jessica Kingsley Publishers, 1999.

Davis & Taylor 1999

Davis, R. & Taylor, B. (1999). Does batterer treatment reduce violence? A synthesis of the literature. Women & Criminal Justice, 10(2), pp. 69-93.

Babcock, Green & Robie 2003

Babcock, Julia, Green, Charles and Robie, Chet (2003). Does batterer treatment work? A meta-analytic review of domestic violence treatment. Clinical Psychology Review, 23, 1023-1053.

Social, Psychological, Criminological and Educational Trials Register (SPECTR)

<http://www.aic.gov.au/campbellcj/studies.html>

PsiTri database of randomized and controlled trials in mental health

<http://www.terkko.helsinki.fi/eu-psi/psitri.htm>

Appendix Two: Coding Protocol (Revision 2.17.04)

Study Level Code Sheet

Use one study level code sheet for each study.

Identifying Information:

1. Study (document) identifier StudyID _____
2. Cross reference document identifier CrossRef1 _____
3. Cross reference document identifier CrossRef2 _____
4. Cross reference document identifier CrossRef3 _____
5. Coder's initials SCoder _____
6. Date coded SDate ____ - ____ - ____

General Study Information:

7. Author Author _____
8. Funder (e.g., NIJ) Funder _____
9. Geographical Location of Study SLocale _____
10. Date range for participant entry into study (preferably when sample pulled):
StartDate ____ - ____ - ____
DoneDate: ____ - ____ - ____
11. Publication Type PubType ____
 1. Book
 2. Book Chapter
 3. Journal (peer reviewed)
 4. Federal Gov't Report
 5. State/Local Gov't Report
 6. Dissertation/Thesis
 7. Unpublished (tech report, conference paper)
12. Number of treatment groups TxGrps ____
13. Number of control groups CgGrps ____
14. Is the same control group used in different contrasts
(1=yes, 0=no, 8=NA) SameCG ____

Treatment-Comparison Level Code Sheet

Use one treatment-comparison level code sheet for each treatment-comparison within a study. For example, if a study has three treatment conditions and each is compared to a single control condition, code the information below separately for each treatment

compared to the single control condition resulting in three treatment-comparison code sheets. Give each treatment-comparison a unique treatment-comparison identifier (TxID), such as 1, 2, 3, etc.

Identifying Information:

15. Study (document) identifier StudyID _____

16. Treatment-comparison identifier TxID _____

Label for this treatment _____

17. Coder's initials TCoder _____

Nature of the Treatment:

18. Type of treatment program (code all that apply, 1=yes, 0=no, 9=cannot tell):

- | | |
|--|---------------|
| (a) Cognitive-behavioral | TxType1 _____ |
| (b) Psychoeducational (including Duluth) | TxType2 _____ |
| (c) Feminist | TxType3 _____ |
| (d) Individual counseling | TxType4 _____ |
| (e) Marital counseling | TxType5 _____ |
| (f) Extensive Monitoring | TxType6 _____ |
| (g) Other _____ | TxType9 _____ |

19. Treatment existed in the community prior to the research study
(1=yes, 0=no, 9=cannot tell) TxExist _____

20. Treatment format (1=yes, 0=no, 9=cannot tell)

1. Group (men only)	TxFrmat1 _____
2. Conjoint group	TxFrmat2 _____
3. Individual (offender)	TxFrmat3 _____
4. Individual (victim)	TxFrmat4 _____
5. Conjoint (couples) individual	TxFrmat5 _____
6. Other _____	TxFrmat8 _____

21. Does the treatment group also receive probation?
(1=yes, 0=no, 9=cannot tell) TxProb _____

22. Average length of probation in weeks (4.3 weeks per month; 88=not applicable; 99=missing) TxProbLn _____

23. How voluntary is the offender's participation? TxVolun _____

1. Nonvoluntary (court mandated)
2. Court-mandated after agreement from offender
3. Completely voluntary
4. Some court-mandated, some voluntary
9. Cannot tell

24. If #4 above, specify the percent court mandated (888 if n/a) TxMand _____

25. Duration of the treatment program in weeks (99 if unknown) TxWeeks ____
26. Number of treatment sessions (99 if unknown, 88 if not applicable) TxNum ____
27. Mean number of sessions attended
(99 if unknown, 88 if not applicable) TxAttend ____
28. Length of a treatment session in hours (99 if unknown, 88 if not applicable) TxHours1 ____
29. Total length of treatment in hours (99 if unknown, 88 if not applicable) TxHours2 ____
30. Sanctions applied for failing to comply with treatment? TxSanc ____
1. Yes, typically
 2. Yes, sporadically
 3. No
 9. Not indicated
31. How many sessions was considered successful completion of treatment?
(99 if unknown, 88 if not applicable) TxNumSuc ____
32. Percentage of treatment sample completing program, as defined by the
researchers. (999 if missing) TxCompl ____
33. Evidence of adherence to the treatment protocol (i.e., fidelity of the treatment
delivery separate from subject compliance; did the treatment providers adhere to
the treatment model; any evidence of fidelity considered, including certification)
TxFidel ____
1. Yes, evidence of treatment fidelity
 2. Evidence of some deviation from treatment model
 3. Evidence of serious deviation from treatment model
 4. No mention of treatment fidelity

Nature of the Control Condition:

34. Who was included in the comparison group? CgVol ____
1. Voluntary treatment or program seekers only
 2. Arrested individuals
 3. Mix of the above individuals
35. What does the comparison group receive? (Code 1=yes, 0=no, 9=not indicated for
each) CgType1 ____
1. No treatment or program
 2. Probation CgType2 ____
 3. Jail or Prison CgType3 ____
 4. Community service CgType4 ____
 5. Some or no treatment (e.g., treatment no-show or drop out) CgType5 ____
 6. Other CgType6 ____
 7. Cannot tell CgType7 ____

36. Duration of comparison group program (in weeks, 99=missing): CgWeeks ____

37. Investigated compensation for control group
(1=yes, 0=no, 9=not indicated) CgComp ____

Methodological Rigor:

38. How were subjects assigned to conditions? TxRandom ____
1. Random (simple)
2. Random (matching pairs)

Quasi-random (alternative cases, alternative blocks of cases)

Other _____

39. Misassignment rate (percentage of cases that violated the random assignment protocol) (999 if missing, 888 if not applicable)
(a) From treatment to control TxMsRte1 ____
(b) From control to treatment TxMsRte2 ____
(c) Total TxMsRte3 ____

40. How did the researchers handle violations to random assignment? TxAnalyz ____
1. Analyzed as assigned
2. Analyzed as treated
3. Both 1 and 2 above (only code effect sizes for 1)
4. Removed cases
8. Not applicable
9. Not indicated

41. Type of quasi-experimental design (nature of comparison group) CgNature ____
01. Historical comparison group
02. Judge or prosecutor did not court-mandate into treatment but gave them an alternative sanction (Specify: _____)
03. Treatment no-shows or drop-outs as the comparison group
04. Eligible domestic violence offenders from an alternative jurisdiction without a court mandated program
05. Matched samples comparison group (sample of non-court mandated domestic offenders drawn from a large pool in a manner designed to produce a group with similar background characteristics to the court mandated group)
06. Domestic violence offenders referred but not accepted into the treatment program
07. Domestic violence offenders not arrested (e.g., Syers and Edleson study)
08. Other _____
88. Not applicable (experimental design)
99. Not indicated

42. Who was included in the experimental group for the comparison? TxCg ____
1. All individuals assigned to treatment (includes rejects, no-shows & drop-outs)
2. All individuals assigned to treatment (excludes rejects)

3. Only those who completed a specified amount of treatment
 4. Only those who completed all treatment
 8. Other _____
43. Did the researchers test for baseline (pre-test) differences?
(1=yes; 0=no) TxDiff1 ____
44. If yes to above, nature of any pretest differences TxDiff2 ____
1. If $n \geq 100$, no significant differences
 2. If $n < 100$, no substantive or significant differences
 3. Minor differences or differences on variables unlikely to be related to offending
 4. Major or important difference
 8. Not applicable
45. Baseline (pretest) differences judged to bias the results in which direction? TxBias ____
1. Positive bias (treatment effect likely to be larger than it really is)
 2. Negative bias (treatment effect likely to be smaller than it really is)
 3. No bias (no differences or differences on variables that should have no effect)
 4. Cannot make a judgment (differences have an uncertain effect)
 8. Not applicable (answered no to question 43)
 9. Cannot tell
46. Analysis of treatment effect statistically adjusted for baseline differences (e.g., logistics regression and ANCOVA that included baseline and background characteristics) (1=yes; 0=no; 9=cannot tell) StatCtrl ____
47. If not an experimental design, was matching used? (1=yes; 0=no; 8=na; 9=cannot tell) MchCtrl ____
48. Which, if any, of the following baseline/background characteristics were matched on or statistically controlled for in the estimate of the treatment effect? (1=yes; 0=no; 9=cannot tell)
- | | |
|---|------------|
| 1. Age | Cov01 ____ |
| 2. Race/ethnicity | Cov02 ____ |
| 3. Employment status | Cov03 ____ |
| 4. Income | Cov04 ____ |
| 5. Prior domestic violence history | Cov05 ____ |
| 6. Prior violent history (general or nondomestic) | Cov06 ____ |
| 7. Prior non-violent history | Cov07 ____ |
| 8. Seriousness of present offense | Cov08 ____ |
| 9. Education | Cov09 ____ |
| 10. Marital status | Cov10 ____ |
| 11. Alcohol or drug use | Cov11 ____ |
| 12. Psychosocial or personality variables (e.g., MMPI, self-esteem) | Cov12 ____ |
| 88. Other _____ | Cov88 ____ |

Sample Characteristics:

Note: These questions apply to the sample characteristics for the subjects included in both the treatment and control conditions for this treatment-comparison contrast. If there are multiple treatment-comparison conditions and data are presented for the study as a whole, use the overall data.

49. Total sample size for this treatment comparison (at start of study) (9999=missing) STotN _____
50. Treatment sample size (at start of study) (9999=missing) STxN _____
51. Control sample size (at start of study) (9999=missing) SCgN _____
52. Sample characteristics (1=yes, 0=no, 9=not indicated)
- (a) nonvoluntary (court referred) SCtRef _____
 - (b) misdemeanor defendants/offenders SMisdem _____
 - (c) misdemeanor and felony defendants/offenders SMisFel _____
 - (d) voluntary SVol _____
 - (e) other _____ SOther _____
53. Treatment sample's disposition (1=yes, 0=no, 9=not indicated)
- (a) post-conviction TPostC _____
 - (b) conditional discharge TCDisch _____
 - (c) pretrial diversion (adjourned in contemplation of dismissal) TPrtial _____
 - (d) military disposition _____ TMilitary _____
 - (e) other _____ TSOther _____
54. Control sample's disposition (1=yes, 0=no, 9=not indicated)
- (a) post-conviction CPostC _____
 - (b) conditional discharge CCDisch _____
 - (c) pretrial diversion (adjourned in contemplation of dismissal) CPrtial _____
 - (d) military disposition _____ CMilitary _____
 - (e) other _____ CSOther _____
55. Sample demographics for treatment comparison sample
- (a) mean age (99 if missing) SMAge _____
 - (b) mean educational level (99 if missing) SMEduc _____
 - (c) % married (999 if missing) SPerM _____
 - (d) % African American (999 if missing) SPerAA _____
 - (e) % Hispanic (999 if missing) SPerHisp _____
 - (f) % employed (999 if missing) SPerEmp _____
 - (g) % with prior arrest (999 if missing) SPriorA _____
56. Was the abuse verified in some form (including conviction)? (2=yes, 1=for some, 0=no, 9=not indicated) SVerify _____
57. Was the sample restricted in anyway (beyond exclusively using heterosexual intimate partners)? (1=yes, 0=no, 9=not indicated) SRstrct _____
58. If yes, indicate nature of restriction (Code 1=yes, 0=no, 8=n/a, 9=missing for each)
- (a) age SRtrct01 _____
 - (b) language SRtrct02 _____
 - (c) geographical area SRtrct03 _____
 - (d) alcohol or drug dependence/abuse SRtrct04 _____

- | | |
|-----------------------------------|---------------|
| (e) mental illness | SRtrct05 ____ |
| (f) criminality | SRtrct06 ____ |
| (g) defendant had to agree | SRtrct07 ____ |
| (h) victim had to agree | SRtrct08 ____ |
| (i) judge had to agree | SRtrct09 ____ |
| (j) prosecutor had to agree | SRtrct10 ____ |
| (k) defense attorney had to agree | SRtrct11 ____ |
| (l) other _____ | SRtrct88 ____ |

59. Do any of the restrictions above result in a sample that is more likely to respond positively to treatment than the general population of domestic violence offenders? (i.e., is the sample "creamed" by eliminating unmotivated or otherwise difficult to treatment offenders?) (1=yes, 0=no, 8=not applicable, 9=cannot tell) SCream ____
60. What proportion of the population of domestic violence offenders in this jurisdiction is the sample? (999 if not indicated) SProp _____

Miscellaneous

61. Does the study examine (i.e., mention in the report) the relationship between treatment attendance (dose) and recidivism (1=yes, 0=no) DoseRel ____
62. If yes to #54, specify the nature of the observed relationship DoseEff ____
1. Negative (higher attendance, less recidivism), and statistically significant
 2. Negative and statistically nonsignificant
 3. Positive (higher attendance, higher recidivism), and statistically significant
 4. Positive and statistically nonsignificant
 5. Statistically nonsignificant, no direction reported
 6. Statistically nonsignificant and correlation equals 0
 8. Not applicable (i.e., answered no to 61)

Outcome (Dependent Variable) Level Code Sheet

Code the information below separately for each dependent variable (outcome) for which an effect size will be coded. Note that time of measurement is on the effect size level code sheet. As such, an outcome measured at multiple time points (rearrest at 6-months, 12-months, and 24-months) should only be coded once using this sheet.

Identifying Information:

63. Study (document) identifier StudyID ____
64. Dependent measure identifier DVID ____
65. Coder's initials DVCoder ____
66. Date coded DVDate ____ - ____ - ____

Dependent Variable Information:

67. Label _____

68. Source of information DVSource ____
1. Official reports (police reports, etc.)
 2. Victim report
 3. Offender self-report
 4. Victim report and offender self-report
 5. Victim report and official report
 6. Victim report, offender self-report and official report
 7. Other _____
69. What is the variable measuring? DVCnstr ____
1. domestic/partner abuse (same partner as initial offense resulting in entry into study)
 2. domestic/partner abuse (same or other partner)
 3. violent crime, excluding domestic/partner abuse
 4. violent crime, includes domestic/partner abuse and other violent offenses (i.e., unspecified person offenses)
 5. drug related offense
 6. property offense
 7. unspecified violation of probation (not specific to 1-5 above)
 8. any type of offense (excluding technical violations)
 9. any type of offense (including technical violations)
 10. other behavioral _____
 11. other nonbehavioral (attitudinal) _____
70. If outcome includes technical violations, was failing to comply with treatment a violation of probation for the experimental conditions? (1=yes, 0=no, 8=not applicable, 9=cannot tell) DVViolate ____
71. For official report measures, indicate the nature of the indicator. (Code 1=yes, 0=no, 8=not applicable, 9=cannot tell for each of the categories below)
1. Official complaint (complaint made to the police, probation officer or judge; may or may not have resulted in an arrest) DVType1 ____
 2. Arrest DVType2 ____
 3. Conviction DVType3 ____
 4. Other _____ DVType4 ____
72. Level of measurement DVLOM ____
1. Dichotomous indicator
 2. Frequency count
 3. Composite scale (semi-continuous)
 4. Other _____

Effect Size Level Coding Sheet

Code this sheet separately for each eligible effect size.

Identifying Information:

73. Study (document) identifier StudyID _____

74. Treatment-Comparison identifier ESID _____
75. Outcome (dependent variable) identifier ESID _____
76. Effect size identifier ESID _____
77. Coder's initials ESCoder _____
78. Date coded ESDate ____ - ____ - ____

Effect Size Related Information:

79. Months from assignment to conditions to point of measurement. (Note: This represents the total time from the start of an individual's involvement in the study to the measurement point for this effect size. A 12-month post assignment to conditions indicator of recidivism would be coded as 12. A 12-month post the end of treatment would be coded 12+length of treatment. Record 99 if information is not available and 88 if not applicable)

- (a) Same for all subjects TimeMon1 _____
- (b) Average (if different across subjects) TimeMon2 _____
- (c) Minimum (if different across subjects) TimeMon3 _____
- (d) Maximum (if different across subjects) TimeMon4 _____

80. Time frame represented by measure in months. (Note: This represent the timeframe for the measure. For example, if arrest data were examined for a 12 month period, then record 12. Similarly, if a survey question asked about abuse during the past 30 days, record 1.)

- (a) Same for all subjects ESTime1 _____
- (b) Average (if different across subjects) ESTime2 _____
- (c) Minimum (if different across subjects) ESTime3 _____
- (d) Maximum (if different across subjects) ESTime4 _____

81. What Is the start of the time frame? StrtTime ____

1. Arrest
2. Conviction and/or Sentence
3. Assignment to conditions
4. Conviction and/or Sentence and/or Assignment to conditions
5. End of treatment for treatment group
6. End of treatment for individual subjects
8. Other _____
9. Cannot tell

82. Direction of effect. (Note: Specify the direction of the effect. Do not leave as missing or this effect size cannot be used.)

ESDirect ____

1. Effect favors experimental (treatment) condition
2. Effect favors control condition
3. Effect favors neither condition (no difference; effect size equals 0)
9. Cannot tell

Effect Size Data—All Effect Sizes:

83. Treatment group sample size ES_TxN _____

84. Control group sample size ES_CgN _____

Effect Size Data—Continuous Type Measures:

85. Treatment group mean ES_TxM _____

86. Control Group mean ES_CgM _____

87. Are the above means adjusted (e.g., ANCOVA adjusted)? (1=yes, 0=no)
ES_MAdj _____

88. Treatment group standard deviation ES_TxSD _____

89. Control group standard deviation ES_CgSD _____

90. Treatment group standard error ES_TxSE _____

91. Comparison group standard error ES_CgSE _____

92. t -value from an independent t -test or square root of F -value from a one-way analysis of variance with one df in the numerator (only two groups) ES_t _____

Effect Size Data—Dichotomous Measures:

93. Treatment group; number of failures (recidivators) ES_TxNf _____

94. Comparison group; number failures (recidivators) ES_CgNf _____

95. Treatment group; proportion failures ES_TxPf _____

96. Comparison group; proportion failures ES_CgPf _____

97. Are the above proportions adjusted for pretest variables?
(1=yes; 0=no) ES_PAdj _____

98. Logged odds-ratio ES_LgOdd _____

99. Standard error of logged odds-ratio ES_SELgO _____

100. Logged odds-ratio adjusted? (e.g., from a logistic regression analysis with other independent variables)
(1=yes; 0=no) ES_OAdj _____

101. Chi-square value with $df = 1$ (2 by 2 contingency table)
ES_ChiSq _____

102. Correlation coefficient (phi) ES_RPhi _____

Effect Size Data—Hand Calculated:

103. Hand calculated d -type effect size

ES_Hand1 _____

104. Hand calculated standard error of the d -type effect size

ES_Hand2 _____